```
FILE 'HOME' ENTERED AT 08:56:37 ON 17 APR 2002

=> file .bioinformatics

COST IN U.S. DOLLARS                                SINCE FILE      TOTAL
                                                       ENTRY    SESSION
FULL ESTIMATED COST                                     0.21       0.21

FILE 'MEDLINE' ENTERED AT 08:56:57 ON 17 APR 2002

FILE 'CAPLUS' ENTERED AT 08:56:57 ON 17 APR 2002
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'BIOBUSINESS' ENTERED AT 08:56:57 ON 17 APR 2002
COPYRIGHT (C) 2002 Biological Abstracts, Inc. (BIOSIS)

FILE 'BIOCOMMERCE' ENTERED AT 08:56:57 ON 17 APR 2002
COPYRIGHT (C) 2002 BioCommerce Data Ltd. Richmond Surrey, United Kingdom. All rights
reserved

FILE 'BIOTECHABS' ACCESS NOT AUTHORIZED

FILE 'COMPENDEX' ENTERED AT 08:56:57 ON 17 APR 2002
COPYRIGHT (C) 2002 ENGINEERING INFORMATION, INC. (EI)

FILE 'COMPUAB' ENTERED AT 08:56:57 ON 17 APR 2002
COPYRIGHT (C) 2002 Cambridge Scientific Abstracts (CSA)

FILE 'COMPUSCIENCE' ENTERED AT 08:56:57 ON 17 APR 2002
COPYRIGHT (c) 2002 FACHINFORMATIONSZENTRUM KARLSRUHE (FIZ KARLSRUHE)

FILE 'EMBASE' ENTERED AT 08:56:57 ON 17 APR 2002
COPYRIGHT (C) 2002 Elsevier Science B.V.  All rights reserved.

FILE 'INSPEC' ENTERED AT 08:56:57 ON 17 APR 2002
Compiled and produced by the IEE in association with FIZ KARLSRUHE
COPYRIGHT 2002 (c) INSTITUTION OF ELECTRICAL ENGINEERS (IEE)

FILE 'JICST-EPLUS' ENTERED AT 08:56:57 ON 17 APR 2002
COPYRIGHT (C) 2002 Japan Science and Technology Corporation (JST)

FILE 'MATH' ENTERED AT 08:56:57 ON 17 APR 2002
COPYRIGHT 2002 (c) FACHINFORMATIONSZENTRUM KARLSRUHE (FIZ Karlsruhe)

FILE 'SCISEARCH' ENTERED AT 08:56:57 ON 17 APR 2002
COPYRIGHT (C) 2002 Institute for Scientific Information (ISI) (R)

=> s (phenotyp? or physical or trait) and database and (sequence or DNA or nucleic acid)

   2 FILES SEARCHED...
   8 FILES SEARCHED...
L1         3681 (PHENOTYP? OR PHYSICAL OR TRAIT) AND DATABASE AND (SEQUENCE OR
                DNA OR NUCLEIC ACID)

=> dup rem l1

DUPLICATE IS NOT AVAILABLE IN 'BIOCOMMERCE'.
ANSWERS FROM THESE FILES WILL BE CONSIDERED UNIQUE
PROCESSING IS APPROXIMATELY  51% COMPLETE FOR L1
```

```
PROCESSING IS APPROXIMATELY  76% COMPLETE FOR L1
PROCESSING COMPLETED FOR L1
L2          1809 DUP REM L1 (1872 DUPLICATES REMOVED)

=> focus l2

FOCUS NOT AVAILABLE IN 'MATH'.
FOCUS NOT AVAILABLE IN 'COMPUSCIENCE'.
PROCESSING COMPLETED FOR L2
ANSWERS FROM NON FOCUS FILES PUT AT END OF ANSWER SET.
L3          1809 FOCUS L2 1-

=> d bib ab 1-150
```

L3 ANSWER 129 OF 1809 SCISEARCH COPYRIGHT 2002 ISI (R)
AN 97:450595 SCISEARCH
GA The Genuine Article (R) Number: XD192
TI Direct classification and selection of superior alleles for crop improvement
AU Sorrells M E (Reprint); Wilson W A
CS CORNELL UNIV, DEPT PLANT BREEDING & BIOMETRY, 252 EMERSON HALL, ITHACA, NY 14853 (Reprint)
CYA USA
SO CROP SCIENCE, (MAY-JUN 1997) Vol. 37, No. 3, pp. 691-697.
Publisher: CROP SCIENCE SOC AMER, 677 S SEGOE ROAD, MADISON, WI 53711.
ISSN: 0011-183X.
DT General Review; Journal
FS AGRI
LA English
REC Reference Count: 46
ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS
AB The use of conventional breeding methods has resulted in consistent crop improvement within the cultivated gene pool by producing genotypes with new combinations of alleles that produce better **phenotypes** than either of the parents (transgressive segregation). Biotechnology has provided new methods to generate, identify, characterize, and manipulate genetic variation. Among these methods, marker assisted selection (MAS) with **DNA** markers has been applied in plant improvement programs. However, MAS is limited by the effort required to generate information about map location and breeding value of genes controlling important **traits**. Comparative genetic analysis across the domesticated grasses facilitates the identification and localization of gene **sequences** controlling specific **traits** of interest. The emerging **databases** of gene **sequences** will allow directed discovery of genes in higher plants and classification of alleles present within breeding germplasm. Identification of the genes controlling a **trait** and knowledge of their **DNA sequence** would facilitate classification of variation in the germplasm pool based on gene fingerprinting or characterization of variation in key **DNA sequences**. Classification of **sequence** variants at a targeted locus would substantially reduce the amount of work required to determine their relative breeding value and lead to the identification of superior alleles. Combining direct allele selection (DAS) with conventional selection, would allow more rapid and precise improvement of populations and breeding lines. Limitations of current technology can be minimized by transfer of genetic information across species, identification of highly variable genes, and focusing on the most important genes and **traits** for the species of interest.

L3 ANSWER 104 OF 1809 CAPLUS COPYRIGHT 2002 ACS
AN 1999:363100 CAPLUS
DN 131:242737
TI Integration of heterogeneous biotechnology **databases**
AU Jones, Simon B.; Franklin, Jack
CS CAB International, Wallingford, OX10 8DE, UK
SO Proc. Int. Chem. Inf. Conf. (1998), 140-148. Editor(s): Collier, Harry. Publisher: Infonortics Ltd., Tetbury, UK.
CODEN: 67SSAV
DT Conference
LA English
AB Researchers in the disciplines related to biotechnol. need to make use of a wide range of information resources in their work. The relatively new discipline of Bioinformatics has assumed considerable importance in providing the infrastructure required for the science. However, the resources available are not generally coherently organized and are consequently not used to max. effect to advance knowledge. The development of a strategy for integrating the information resources is described and the technol. being used to provide a practical implementation of demonstration to a target user group is outlined. This project brings together a group of **databases** available as both a **phys**. co-located collection on one or more established hosts, and as a looser federation of **databases** accessible over the Internet. Sophisticated data parsing mechanisms enable links between data of similar and dissimilar types so that researchers can follow through a research story. The project **databases** encompass genetic and protein **sequences**, resource collections, ongoing projects and com. exploitation, as well as access to primary and secondary published literature sources, including three major secondary **databases**. A review of user reactions to the demonstrations and of the processes leading to an exploitation plan for the results is also included.
RE.CNT 10 THERE ARE 10 CITED REFERENCES AVAILABLE FOR THIS RECORD
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 99 OF 1809 CAPLUS COPYRIGHT 2002 ACS
AN 2000:513433 CAPLUS
DN 133:117580
TI Analysis of patterns of gene expression in plant tissues as an indicator of heterosis in breeding for agronomic **traits**
IN Bowen, Ben; Guo, Mei; Smith, Oscar
PA Pioneer Hi-Bred International, Inc., USA
SO PCT Int. Appl., 74 pp.
CODEN: PIXXD2
DT Patent
LA English
FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	WO 2000042838	A2	20000727	WO 2000-US1422	20000119
	WO 2000042838	A3	20010329		
	W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
	RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
	EP 1143787	A2	20011017	EP 2000-904457	20000119
	R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO				
PRAI	US 1999-116617P	P	19990121		
	US 1999-166368P	P	19991117		
	WO 2000-US1422	W	20000119		

AB Methods of correlating mol. profile information and heterosis are provided. Selection for dominant, additive, or under/overdominant markers provides for improved heterosis. Selection for the no. of expression products in an expression profile provides for improved heterosis. Methods of identifying and cloning **nucleic acids** linked to heterotic **traits** are provided. Methods of identifying parentage by consideration of expression profiles are provided. The invention finds that heterosis is a function of the no. of genes expressed at their optimum levels. The use of large-scale anal. of patterns of gene expression to study heterosis is demonstrated.

L3 ANSWER 96 OF 1809 INSPEC COPYRIGHT 2002 IEE
AN 1995:5134359 INSPEC DN C9601-7330-199
TI Integrated access to heterogeneous data from NCBI.
AU Ostell, J.M.
SO IEEE Engineering in Medicine and Biology Magazine (Nov.-Dec. 1995) vol.14, no.6, p.730-6. 2 refs.
Price: CCCC 0739-5175/95/$4.00
CODEN: IEMBDE ISSN: 0739-5175
DT Journal
TC General Review; Practical
CY United States
LA English
AB The National Center for Biotechnology Information (NCBI) was created by Congress in 1988 to provide a stable government entity with a leadership role in coping with the information issues associated with this field. The NCBI Software Toolkit permits software tools to be developed in a heterogeneous environment. The use of abstract syntax notation (ASN.I) allows one to specify and exchange data across systems, and to reach biologists in whatever system they choose to work. However, the relevant data are still accumulated by different groups with different data models, different quality standards, in different subject domains, and with different time courses. NCBI has designed a data model to define a number of key data elements for molecular biology, including bibliographic data, **nucleic acid sequence**, protein **sequence**, genetic and **physical** maps, and the information about them. The model was constructed in as much detail as possible to accommodate the data contained in the heterogeneous sources, while still maintaining a common model.

L3 ANSWER 68 OF 1809 CAPLUS COPYRIGHT 2002 ACS
AN 2000:894855 CAPLUS
DN 135:205934
TI An essay on individual **sequence** variation in expressed **sequence** tags (ESTs)
AU Reich, Jens; Brett, David; Hanke, Jens
CS Max Delbruck Center of Molecular Medicine Charite Medical Faculty, Humboldt University of Berlin, Germany
SO Genomics and Proteomics: Functional and Computational Aspects, [Proceedings of the International Symposium on Genomics and Proteomics: Functional and Computational Aspects], Heidelberg, Germany, Oct. 4-7, 1998 (2000), Meeting Date 1998, 83-94. Editor(s): Suhai, Sandor. Publisher: Kluwer Academic/Plenum Publishers, New York, N. Y. CODEN: 69ATP7
DT Conference; General Review
LA English
AB A review with 31 refs. Expressed **sequence** tags (ESTs) are short **sequence** segments (usually up to 500 nt long) obtained by reverse-transcription into cDNA clones from mRNA prepns. of a cell or tissue in a specified functional or developmental stage. The main application of ESTs is anal. of which cell type expresses which gene. Also, gene expression in a given cell type can be studied in different functional and developmental stages. ESTs can be used to study pathol. conditions of a cell; for example, the expression pattern of tumor cells may be compared to that of normal tissue. ESTs can help in the search for candidate genes responsible for certain **traits** through homol. searches on other species in genomic **databases**. STSs (**sequence** tag sites), ESTs that have been mapped to their chromosomal location, serve as valuable markers of coding **sequences** within the genome. Each EST stems from a unique specimen of a particular species, and may therefore reveal individual variation of the expressed genomic information: gene polymorphism, gene defects, mutations, or splice variants. Clusters of aligned ESTs, assembled into gene segments, are available in the public domain, permitting identification of single nucleotide polymorphisms (SNPs), deletions, insertions, and splicing variants.
RE.CNT 31 THERE ARE 31 CITED REFERENCES AVAILABLE FOR THIS RECORD
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 24 OF 1809 CAPLUS COPYRIGHT 2002 ACS
AN 2000:606782 CAPLUS
DN 133:188871
TI Processes, apparatus and compositions for fingerprinting **nucleic acids** by k-tuple analysis
IN Lopez-Nieto, Carlos Eduardo; Nigam, Sanjay Kumar
PA Brigham and Women's Hospital, USA
SO U.S., 47 pp., Cont.-in-part of U.S. Ser. No. 242,887, abandoned.
CODEN: USXXAM
DT Patent
LA English
FAN.CNT 2

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	US 6110667	A	20000829	US 1996-522384	19961115
	WO 9531574	A1	19951123	WO 1995-US6032	19950512
	W: AU, CA, JP, US				
	RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
PRAI	US 1994-242887	B2	19940516		
	WO 1995-US6032	W	19950512		

AB A sample contg. nucleotide **sequences** is characterized or "fingerprinted" by contacting it, or portions thereof, with of plurality of primer pairs; detecting nucleotide regions delineated by subsequences hydridized with each primer pair; detg. a **phys.** characteristic of each such detected region; and indexing such **phys.** characteristics as a function of the primer pair that hydridized to the delineating subsequences of the corresponding region. An app. for fingerprinting a sample contg. nucleotide **sequences** includes functionality for carrying out such a process. A kit of primers for use in characterizing, or fingerprinting, a sample comprising one or more nucleotide **sequences** is produced by a process identifying a first set of primers that hybridize with relative high frequency to resp. antisense subsequences putatively present in the sample (or complementary subsequences thereof); identifying a second set of primers that hybridize with relative low frequency to resp. sense subsequences putatively present in the sample (or complementary subsequences thereof); and, providing as the kit of primers at least selected primers common to both the first and second sets. Design of primers and their use in fingerprinting patterns of G protein-coupled receptor gene expression are described.

RE.CNT 9 THERE ARE 9 CITED REFERENCES AVAILABLE FOR THIS RECORD
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3

L3 ANSWER 19 OF 1809 CAPLUS COPYRIGHT 2002 ACS
AN 2001:320104 CAPLUS
DN 134:336663
TI Datamining of biosequence **databases** for genes affecting a given **phenotype**
IN Bulla, Lee A., Jr.; Candas, Mehmet
PA USA
SO PCT Int. Appl., 73 pp.
CODEN: PIXXD2
DT Patent
LA English
FAN.CNT 1

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI WO 2001031011	A2	20010503	WO 2000-US29445	20001025
WO 2001031011	A3	20020117		

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRAI US 1999-161527P P 19991026

AB The present invention provides a system, method and app. for targeting gene **sequences** having one or more **phenotypic** characteristics using a computer. One or more **phenotypic** characteristics are selected. A gene **sequence** is then selected that is known to have the selected **phenotypic** characteristics. In addn., one or more **databases** contg. cataloged gene **sequences** are selected. The selected gene **sequence** is compared to the cataloged gene **sequences**, and any cataloged gene **sequences** that contain a portion of the selected gene **sequence** are extd. The selected gene **sequence** is aligned to each portion of the extd. gene **sequence** and the extd. gene **sequences** are prioritized based on the alignment of the selected gene **sequence**. At least one of the prioritized gene **sequences** is selected based on one or more **phenotypic** criteria. Finally, one or more degenerate primers are designed to target the selected-prioritized gene **sequences**.

L3 ANSWER 17 OF 1809 CAPLUS COPYRIGHT 2002 ACS
AN 1998:805789 CAPLUS
DN 130:233207
TI An introduction to Internet resources for the molecular and genetic analysis of the lipases
AU Reue, Karen
CS Department of Medicine, and West Los Angeles VA Medical Center, University of California, Los Angeles, CA, USA
SO Methods Mol. Biol. (Totowa, N. J.) (1999), 109(Lipase and Phospholipase Protocols), 309-328
CODEN: MMBIED; ISSN: 1064-3745
PB Humana Press Inc.
DT Journal
LA English
AB The goal of this chapter is to provide an overview of some of the established online **databases** that provide access to **nucleic acid** and protein **sequences**, **DNA** reagents, gene mapping data, mutations and **phenotype** information, and available animal models. This chapter lists the addresses and key features of several **databases** that are of general interest and utility, and provides guidelines and practical tips for using selected web sites.
RE.CNT 16 THERE ARE 16 CITED REFERENCES AVAILABLE FOR THIS RECORD
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 16 OF 1809 CAPLUS COPYRIGHT 2002 ACS
AN 1992:254879 CAPLUS
DN 116:254879
TI Non-**sequence databases** for biological activity and physicochemical properties
AU Jone, C. S.; Tsugita, A.; Satake, K.; Okibayashi, F.; Imai, K.; Yagi, T.; Takahashi, K.; Yeh, L. S.
CS Res. Inst. Biosci., Sci. Univ. Tokyo, Noda, 278, Japan
SO Protein Sequences Data Anal. (1991), 4(6), 367-74
CODEN: PSDAE6; ISSN: 0931-9506
DT Journal
LA English
AB A biol. activity **database** and physicochem. property **database** are described. They are intended to complement the protein **sequence database** of PIR-International. The Biol. Activity **Database** and the Physicochem. Property **Database** contain information regarding the biol. activity and the physicochem. properties of proteins, resp. In addn., they also provide information about wild-type mols. with which information concerning variant mols. may be compared. Data on artificial variant mols. are stored in the Artificial Variant **Database** which is described sep.

L3 ANSWER 4 OF 1809 CAPLUS COPYRIGHT 2002 ACS
AN 1999:490060 CAPLUS
DN 131:285965
TI Using public **databases**
AU Haines, Jonathan L.
CS Department of Molecular Physiology and Biophysics Program in Human, Vanderbilt University School of Medicine, Nashville, TN, USA
SO Approaches Gene Mapp. Complex Hum. Dis. (1998), 335-349. Editor(s): Haines, Jonathan L.; Pericak-Vance, Margaret A. Publisher: Wiley, New York, N. Y.
CODEN: 67XVAV
DT Conference
LA English
AB There are a large no. of public **databases** of genetic data useful for mapping genes involved in complex **traits**. Although most **databases** of human data are organized around methodologies, other **databases** have attempted to integrate these diverse data types using biol. relevant groupings. These sites are constantly evolving as technol. and experience, and the sheer vol. of data, allow more and better ways of organizing and accessing information. Several useful resources in the Internet are given covering the access and use of genetic marker **databases**, **phys**. mapping **databases**, **sequence databases** and model organism **databases**
.
R

L3 ANSWER 130 OF 1809 MEDLINE
AN 2000063203 MEDLINE
DN 20063203 PubMed ID: 10592169
TI **Database** resources of the National Center for Biotechnology Information.
AU Wheeler D L; Chappey C; Lash A E; Leipe D D; Madden T L; Schuler G D; Tatusova T A; Rapp B A
CS National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Building 38A, 8600 Rockville Pike, Bethesda, MD 20894, USA.
SO NUCLEIC ACIDS RESEARCH, (2000 Jan 1) 28 (1) 10-4.
Journal code: O8L; 0411011. ISSN: 0305-1048.
CY ENGLAND: United Kingdom
DT Journal; Article; (JOURNAL ARTICLE)
LA English
FS Priority Journals
EM 200002
ED Entered STN: 20000314
Last Updated on STN: 20000314
Entered Medline: 20000225
AB In addition to maintaining the GenBank(R) **nucleic acid sequence database**, the National Center for Biotechnology Information (NCBI) provides data analysis and retrieval and resources that operate on the data in GenBank and a variety of other biological data made available through NCBI's Web site. NCBI data retrieval resources include Entrez, PubMed, LocusLink and the Taxonomy Browser. Data analysis resources include BLAST, Electronic PCR, OrfFinder, RefSeq, UniGene, **Database** of Single Nucleotide Polymorphisms (dbSNP), Human Genome Sequencing pages, GeneMap'99, Davis Human-Mouse Homology Map, Cancer Chromosome Aberration Project (CCAP) pages, Entrez Genomes, Clusters of Orthologous Groups (COGs) **database**, Retroviral Genotyping Tools, Cancer Genome Anatomy Project (CGAP) pages, SAGEmap, Online Mendelian Inheritance in Man (OMIM) and the Molecular Modeling **Database** (MMDB). Augmenting many of the Web applications are custom implementations of the BLAST program optimized to search specialized data sets. All of the resources can be accessed through the NCBI home page at: http://www.ncbi.nlm.nih. gov